

R E M A R K S

The Office Action Summary page states in item 6 that claims 1-28, 31-34, 36-40, 43-51, and 57-65 are rejected. This does not correspond either to item 7, which states that claims 20-24 and 27, 28 are objected to, or to the Detailed Action.

Applicants assume that the Detailed Action is controlling and that the Summary page is in error.

The Detailed Action states that claims 1, 2, 12-10, 26, 31-34, 36-40, 43-51, and 57-65 are rejected under 35 USC 103 over Applicants' prior art in view of an article by Bothe, titled "Audio to Audio-video Speech Conversion with the Help of Phonetic Knowledge Integration, IEEE, Jan. 1997. Applicants respectfully traverse.

At page 5 of the Office action, lines 6-10, the Examiner presents what the Examiner considers to be applicants' admitted prior art.

This assertion is found -- verbatim -- in the previous Office action, and in their response (at page 12) applicants traversed the assertion as being inaccurate.

In the instant Office action, in the "Response to Arguments" section, the Examiner points to the first three paragraphs of said page 12 and asserts that applicant "is arguing the newly amended claim language" (page 7, item 7, second sentence), and that "examiner notes that the claim language pertains to text and FAP information, not a TTS stream" (page 7, item 7, third sentence).

Respectfully, the Examiner's position is defective, because it misstates the facts. In explaining the Examiner's flawed understanding of applicants' admitted prior art, no reference to any claim language was made in the first three paragraphs of page 12 in the previous response -- contrary to the Examiner's assertion.

More particularly, applicants clearly explained:

Specifically, applicants admitted that a face model is applied, *but not that the face model is part of the TTS stream*. Likewise, applicants admitted that FAPs are applied to a compositor (through the FRM), *but not that the FAPs are part of the TTS stream*. The Examiner's "along" connective word, however, can be viewed as including a teaching of a TTS stream with embedded FAPs; but as indicated above, that is **not** admitted by applicants to be in the prior art.

Both the italic and the bold highlights in the original, but the Examiner either missed this explanation, or chose to ignore it. There is no rebuttal of the applicants' characterization

and, indeed, there is no reference at all to this explanation in the Examiner's "Response to Arguments."

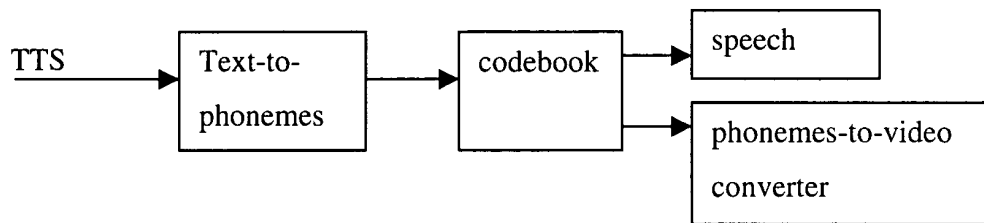
As for the teachings of Bothe, at page 4, lines 13-16, Examiner asserts that:

Bothe (Audio to...) however, teaches the derivation of phonemes from a TTS stream (using a transform) to influence the animation of the phoneme (Fig. 1, pp 1632, second column; pp 16:1. 5, second column - page 1637, end of second column before 'Conclusion and Future Work).

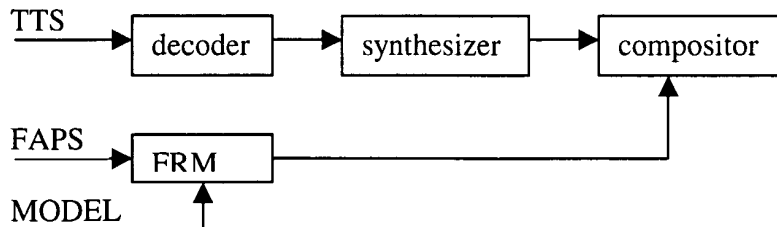
Applicants described the Bothe teaching similarly with the following sentence:

What the Bothe reference teaches is merely the creation of a string of phonemes from applied text and the conversion of the phonemes to video through a conversion that employs a codebook.

Thus, there appears to be substantial agreement between the Examiner and applicants as to what Bothe teaches. Viewed pictorially, Bothe teaches



In contradistinction, the admitted prior art (and it is hoped that there is no dispute here, since it is a copy of FIG. 1) is:



The Examiner makes the conclusory statement that

it would have been obvious to one of ordinary skill in the art of speech animation to modify the teachings of Applicant's Admitted Prior Art, as explained above, with phoneme base animation information because it would advantageously produce an illusion of a mouth moving to generate the sounds heard by synchronizing audio playback of a phoneme with display of a mouth shape corresponding to the phoneme, for application such as video telephones and lip reading training.

Applicants substantively disagree. Moreover, applicants respectfully submit that a proper assessment must include a threshold determination as to whether there is motivation to combine references, and then, there must be an analysis of what would result from such combining.

Applying this approach to the case at hand, the following facts are believed indisputable:

1. The prior art uses FAP information to develop a video. The Bothe reference uses phonemes to develop a video. FAP information is NOT phoneme information.
2. Neither the admitted prior art nor Bothe teach the notion of converting phonemes to FAPs.
3. There is nothing in the admitted prior art to suggest that the method known in the admitted prior art is somehow deficient in creating the "illusion of a mouth moving to generate the sounds heard by synchronizing audio playback of a phoneme with display of a mouth shape corresponding to the phoneme".
4. There also nothing in Bothe to suggest that the method known in the admitted prior art is somehow deficient in creating the aforementioned "illusion," or is in need of improvement.
5. The goal of producing the aforementioned "illusion" is achieved in the Bothe arrangement, but as indicated above, in a different manner. There is no indication that it is superior to the method of the admitted prior art.

For these reasons, applicants respectfully submit that there is no motivation to combine the references. Moreover, applicants respectfully submit that it is not at all clear that a combination of the two references (without hindsight of applicants' claims) would result in a system/method as claimed by applicants. Numerous other possibilities exist.

Turning to the specifics:

As for claim 1, the Examiner asserts that the combination

teaches a decoder responsive to an input signal comprising text and FAP information (applicant's admitted prior art), generating additional [sic] FAP information from phoneme information to add to FAP information....

Applicants disagree. Respectfully, the Examiner's assertion clearly demonstrates the total hindsight of the Examiner's approach. First, there is no suggestion anywhere that the admitted prior art approach is in need of improving. Second, there is no suggestion in

Bothe that its approach is in need of improving. Third, even if one were told to combine the two references, it is not clear -- without hindsight of this invention -- how to combine them, because they both accomplish the same thing by in different ways. Fourth, since there no teaching for converting phonemes to FAPs, there is no reason to believe that -- again without hindsight -- an artisan choose to invent this approach for augmenting, and thus improving, the prior art. Fifth, over and above the reasons already mentioned, the notion of having (a) having a signal that includes FAP information, (b) accepting provided FAP information, (c) and generating additional FAP information is totally novel.

Viewed another way, the Examiner's position would have been stronger if Bothe employed information for generating video that includes phonemes, and then undertook to create some additional information (phonemes, visemes, or any other information), but that is NOT what Bothe teaches.

In short, contrary to the Examiner's characterization of applicants' remarks as "pure conjecture and speculation" (Detailed Action, page 7, item 7, line8), applicant submitted numerous, cogent, reasons that demonstrate that claim 1 is not obvious in view of the admitted prior art combined with the Bothe reference. Claims that depend on claims 1 are also not obvious.

As for claim 2, the Examiner's comment following the rejection of claim 2 seems to be more related to claim 1 than to claim 2. Claim 2 is a method claim that specifies two steps that generate signals. The first step -- that of generating a signal that includes signals for generating a sound -- is admittedly taught by Bothe, as well as by the admitted prior art. The second step, however, that of generating a stream of commands comprising FAP information that excludes viseme information, is clearly not taught or suggested by either the admitted prior art or the Bothe reference. Looking at page 2 of applicants' specification one can find the statement that "a viseme is a visual version of a phoneme." Since Bothe generates a signal for creating a video *from* phonemes, which effectively is saying that Bothe generates visemes, it follows that Bothe generates precisely what claim 2 prohibits. Since Bothe teaches the opposite of what claim 2 specifies, it follows that claim 2 is not obvious in view of the admitted prior art in view of the Bothe reference.

It is noted that, significantly, the Examiner has not explained how he views the admitted prior art and the teachings of Bothe being combined to teach or suggest the above-discussed second step (which is contrary to, and teaches away from, Bothe).

Similarly in connection with claim 12, the claim specifies a decoder that is responsive to an input signal comprising "signals representing audio and embedded video synthesis command signals" (emphasis supplied). The decoder of the admitted prior art is **not** responsive to a signal that is representing audio and embedded video synthesis command signals. It is responsive only to a TTS signal. The decoder of Bothe is also **not** responsive to a signal that is representing audio and embedded video synthesis command signals. It is responsive to a text file and/or to a keyboard. Since neither of the admitted prior art nor the Bothe reference have such a signal, it follows that whatever decoder is found in either the admitted prior art or in Bothe is NOT one that operates as defined in claim 12, to wit,

a decoder ... that separates the command signals from the signals representing audio to develop an audio signal stream and a video synthesis command signals stream,

Hence, claim 12 is not obvious over the admitted prior art in view of Bothe.

Correspondingly, all claims that depend on claim 12 are believed to also be not obvious over the admitted prior art in view of Bothe.

It is noted that the Examiner has provided no justification for asserting obviousness in light of claim 12's clear definition of the composition of the signal applied to the decoder, and the actions performed in the decoder in response to this signal.

The above argument relative to the composition of the input signal also applies to independent claims 31, 43, and 57, and the claims that depend thereon.

Claims 20-25, 27-30, 41, 42, 52-56, and 66-70 were objected to. A number of these claims are amended herein to place all of the objected-to claims in condition for allowance.

In view of the above amendments and remarks, applicants respectfully submit that all of the Examiner's rejections and objections have been overcome. Reconsideration and allowance are respectfully solicited.

It is recognized that this is a response after FINAL. Accordingly, if the Examiner feels that the case is NOT in condition for allowance (in spite of applicants' arguments

above), the Examiner is respectfully requested to call the undersigned. In any event, in the Advisory Action that might follow, the Examiner is respectfully requested to indicate whether the objected-to claims would be allowable if the instant amendment were to be admitted.

Dated: 4/19/04

Respectfully,
Mark Beutnagel
Ariel Fischer
Joern Ostermann
Yao Wang

By 
Henry T. Brendzel
Reg. No. 26,844
Phone (973) 467-2025
Fax (973) 467-6589
email brendzel@home.com